Government Information Quarterly 29 (2012) 470–479

Contents lists available at SciVerse ScienceDirect

Government Information Quarterly

journal homepage: www.elsevier.com/locate/govinf

Opinion mining in social media: Modeling, simulating, and forecasting political opinions in the web

Pawel Sobkowicz [a], Michael Kaschesky [a,*], Guillaume Bouchard [b]

[a] Bern University of Applied Sciences, E-Government Unit, Bern, Switzerland
[b] Xerox Research Center Europe, Machine Learning Group, Meylan, France

ARTICLE INFO

Available online 26 July 2012

Keywords:
Management
Measurement
Design
Experimentation
Opinion mining
Social media
Policy modeling

ABSTRACT

Affordable and ubiquitous online communications (social media) provide the means for flows of ideas and opinions and play an increasing role for the transformation and cohesion of society – yet little is understood about how online opinions emerge, diffuse, and gain momentum. To address this problem, an opinion formation framework based on content analysis of social media and sociophysical system modeling is proposed. Based on prior research and own projects, three building blocks of online opinion tracking and simulation are described: (1) automated topic, emotion and opinion detection in real-time, (2) information flow modeling and agent-based simulation, and (3) modeling of opinion networks, including special social and psychological circumstances, such as the influence of emotions, media and leaders, changing social networks etc. Finally, three application scenarios are presented to illustrate the framework and motivate further research.

1. Introduction

The goal of opinion research is to identify emerging societal trends based on views, dispositions, moods, attitudes and expectations of stakeholder groups or the general public. One major application of opinion research is the area of policymaking in order to better anticipate likely impacts of policy measures and better communicate expected benefits and consequences. Models of opinion formation based on real-world online communication enable the simulation and prediction of the evolution of communication patterns on a specific policy issue within a region or cross-regionally for global comparison.

The democratization of web publishing has led to the explosion of the number of opinions expressed over the internet. At the same time, citizens are becoming more actively engaged in policy issues, more empowered, and more demanding in their relations with traditional institutions while political clubs, organizations, and editorials experience falling memberships (Inglehart & Welzel, 2005).

Research on the blogosphere identifies a ‘hunger’ for and reliance upon peer advice and recommendations found online and this information hunger is strongly evident in the political sphere. For example, through a large-scale survey, researchers were able to infer the motivations of over 60 million U.S. citizens who gathered online information about the 2006 elections and exchanged their views (Rainie & Horrigan, 2007). For one third of these citizens, the motivation to engage online was to get perspectives from inside their community, while another third was motivated by getting perspectives from outside their communities. Another third was motivated by other citizens' endorsements or ratings. The political sphere appears particularly suited for investigating opinion-formation in the blogosphere, because “blogging as democratic practice” is inherently linked to the broader policy processes (Griffiths, 2004).

Affordable and ubiquitous information and communication technologies (ICT) promote the exchange of ideas and opinion across borders. Driving the structural transformation are information flows connecting individual ideas and opinions with others thereby creating the networked society (Castells, 1996). Arguably, the ICT-enabled flows of ideas and opinions play a fundamental role for the transformation and cohesion of the information society – yet little is understood about how online opinions emerge, diffuse, and gain momentum (Christakis & Fowle, 2009). At the same time, the internet provides large amounts of data from online communities making it possible to observe and study social interactions online ‘in situ’. One might treat the internet community as a huge social and psychological laboratory (Skitka & Sargis, 2006). Hence, in this work we follow an overarching research question:

> *In what ways can online content from various social networking resources be exploited to inform decision makers about constituent opinions, emerging trends, and on the feasibility and potential impacts of policy initiatives?*

* Corresponding author.
E-mail addresses: pawelsobko@gmail.com (P. Sobkowicz), ksm1@bfh.ch (M. Kaschesky), guillaume.bouchard@xrce.xerox.com (G. Bouchard).

0740-624X/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.giq.2012.06.005

In the following, we address this question by presenting building blocks for opinion mining, simulation, and understanding of trends:

- *Social media content analysis*: a large set of online forums, blogs or other publicly available text streams are tracked and analyzed. Text understanding algorithms extract semantic information related to the topics targeted by the decision maker. In particular, the social network of individuals expressing their opinion online is reconstructed and for every analyzed text, the main subtopics are identified, as well as the associated sentiment (positive/negative opinions);
- *Opinion formation modeling, simulation and prediction*: An opinion diffusion model is estimated on the extracted data to recover the graph of influence and model current and future opinions' trends. Every opinion is represented by a concept (or sub-topic) and a diffusion rate, and individuals are represented by interests, influence and disposition of being influenced;

The next two sections describe these components in more details. Section 4 then presents application scenarios and real-life examples that naturally fit to the proposed opinion mining framework. Finally, we discuss the practical implementation choices and future research directions.

2. Social media content analysis

Topic and opinion detection in online content facilitates the identification of emerging societal trends and analysis of public reactions to policies. The next step beyond current web search is to rank information entities of varying type, complexity, and structure, rather than document-only (e.g. web pages). Being able to retrieve specific entities rather than whole documents allows building innovative applications for topic and opinion detection (e.g. extracting comments). These possibilities are made possible due to the proliferation of Semantic Web standards and methods, rise of machine learning methods in natural language processing, availability of datasets for machine learning algorithms to be trained on, and the spread of review-aggregation websites and user-rated content. Topic and opinion detection provides a fast and reliable way of transforming a set of unlabeled documents into a well-structured knowledgebase. There are two approaches, which currently develop rather unrelated to each other:

- *Natural language processing (NLP)*: implicit representation of meaning, based on a vector representation of texts and meaning, which enables the definition of similarities between texts and degrees of positive or negative opinions. The outcome of such models is accurate but difficult to interpret.
- *Semantic web approaches (SW)*: explicit representation of the domain based on semantic annotations that map a text to the domain ontology via keywords or tags. There are few large scale examples of efficient reasoning based on this approach.

Today, there are few hybrid systems combining the strengths of both approaches. We present an approach based on a robust method using the implicit representation of meaning (NLP) and extending it using light-weight ontologies for the Semantic Web (SW) to improve performance and allow more fine grained analysis of opinions.

2.1. Opinion detection

The focus is on the automatic identification and extraction of opinions from text and multimedia (Chesley et al., 2006). Motivation for this component is based on providing support for decision makers to automatically track attitudes on certain topics in online media and user generated content (Lin, Wilson, Wiebe, & Hauptmann, 2006). For example, opinion detection has been proposed as a key enabling technology in eRulemaking, allowing the automatic analysis of the opinions that people submit about pending policy or government-regulation proposals (Allen et al., 2005; Kwon, Shulman, & Hovy, 2006; Shulman, Hovy, Callan, & Zavestoski, 2006).

The goal of opinion mining is to create a knowledgebase containing online opinions in a more structured and explicit form. The data is processed by a NLP engine based on a syntax analyzer and machine learning techniques that detect which part of the sentence correspond to the expression of an opinion, and on which specific topic. For each text, the identified opinion is represented as a list of pairs (rhetorical concept, keyword) mentioned in the text. The rhetorical concept is defined a priori by linguists. To start with, the vocabulary will be simplified into four categories, such as 'positive opinion', 'neutral opinion', 'negative opinion' and 'information' (e.g. fact-like information such as quality news).

2.2. Sentiment analysis

Sentiment analysis combines the deliberative with the emotional part (opinions or attitudes with emotions about them). Similarly to the analysis of attitudes, computer-based recognition of emotions requires advanced analytical tools. While lexicon-based solutions provide some level of detecting basic emotions (e.g. by selecting agreeing, confirmative words or detecting swearwords and curses), they fall short of human recognition by the readers, because they often fail to recognize the more subtle forms of expressing emotions: humor, sarcasm, irony, provocation.

Using various categorization algorithms, emotion analysis has already been subject of research (Allen et al., 2005; DeSteno, Petty, Rucker, Wegener, & Braverman, 2004; Prabowo & Thelwall, 2009; Theunis, Küster, Tsankova, & Kappas, 2010). These approaches have proven to be effective analyses of internet-based communities (Chau & Xu, 2006; Chmiel et al., 2011b; Derks, Fischer, & Bos, 2008; Mitrović, Paltoglou, & Tadić, 2010; Thelwall, Wilkinson, & Uppal, 2010). There are also works devoted to modeling of such communication networks and comparisons of the agent-based models and observations (Chmiel et al., 2011a; Ding & Liu, 2010; Schweitzer & Garcia, 2010; Sobkowicz & Sobkowicz, 2010). One of the ways to achieve better agreement between automatic recognition systems and human categorization of emotions is to include capacity to 'remember' evaluation values given to previous expressions attributed to the same anonymized author (*if* such authorship can be identified, which is the case of many social media environments).

It is important to combine the opinion and emotion analyses in a single view. This enables to transcend simplifications typical for sociophysics-based models of opinion change (Sobkowicz, 2009a). For example, our research on discussion fora (Sobkowicz & Sobkowicz, 2010, 2011) shows remarkable stability of individual political opinions of discussion participants. On the other hand, the emotional state is quite flexible. Emotion may dynamically change as a result of a single event: message read, conversation, news item. On the other hand some deeply ingrained opinions (e.g. political affiliation) are impervious to external influences. This phenomenon is easy to understand in the framework of normative/informative processing (Wood, 2000). Within this framework, normative attitudes related to a person's position within the in-group are highly stable; those related to internal coherence of views are moderately stable. In both situations new information is processed in a biased way. On the other end, informative processing, with a balanced approach and flexibility of opinions is associated with information related primarily to specific issues. The more emotional the discussions are, the less chance there is of actually changing participant opinions. With this in mind, we point out the role of moderators and structural properties of internet social media in deciding the deliberative vs. emotional reactions of the users.

3. Opinion formation modeling

Recent years have brought significant interest in interdisciplinary studies, combining tools and methods known from physics with social analyses. These studies are often referred to as sociophysics, and range from purely numerical studies of economic trends to descriptions of social activities. Among the latter, a significant role is played by computational models of opinion formation. Such models often use agent-based simulations. Within a simplified framework, focusing on a few selected aspects of social activities (such as communication network, susceptibility to influences, contrariness etc.), it is possible to derive general trends of behavior of large societal groups, starting from individual perspectives (similar to statistical, kinetic theory of matter).

One of the major problems with 'social physics' or sociophysics research on opinion modeling is the lack of validation based on real-life examples and data. Recent works reiterate the need for a real-life validation by emphasizing real-life evidence over conceptual models and theory (Moss & Edmonds, 2005) as well as prediction and explanation based on real data for opinion modeling and observation (Epstein, 2008). There exists a significant gap between the social and psychological literature devoted to attitude change and models based on physical analogies.

The social and psychological literatures indicate that there are many interacting factors persuading people to change attitudes (for general reviews and accounts of the developments in sociological understanding of these phenomena see, e.g. Cialdini, Levy, Herman, Kozlowski, & Petty, 1976; Crano & Prislin, 2006; Eaton & Visser, 2008; Petty, Wegener, & Fabrigar, 1997; Price, Cappella, & Nir, 2002; Wood, 2000). Moreover, it is clear from field studies that the reactions depend on social context and situation for example (face-to-face contact, or the propensity for deliberation, see Cialdini et al., 1976; Friedkin & Johnsen, 1999; Gastil, Black, & Moscovitz, 2008; Kenny, 1994; Price et al., 2002; Wojcieszak et al., 2010).

On the other hand, the models introduced by the physics community are usually simplifying the psychological description, to allow more rigorous mathematical treatment (for reviews and comments see Castellano, Fortunato, & Loreto, 2009; Sobkowicz, 2009a). The list of factors influencing human attitudes, identified in the psychological literature is long (Petty & Wegener, 1998). We note that the relative importance of these factors is by no means established, and it may differ from one social situation to another. As a result, it is extremely difficult to map the complexity of human behavior into simple, agent-based computer simulations.

While individual sociophysical works are based on rather simple 'models' of human behavior, there are many variants of such approaches, for example the bounded confidence models (Deffuant, Neau, Amblard, & Weisbuch, 2000; Hegselmann & Krause, 2002), the social impact theory (Hołyst, Kacperski, & Schweitzer, 2001; Kacperski & Hołyst, 1999; Nowak & Lewenstein, 1996; Nowak, Szamrej, & Latané, 1990), or the Sznajd local influence model (Sznajd-Weron & Sznajd, 2000). We note that there are also works extending beyond the simple analogies from physics, taking into account dynamical nature of social interactions, different roles played by individual persons, such as leadership or authority positions (Banisch, Araujo, & Louçã, 2010; Huang, Tzou, & Sun, 2011; Sobkowicz, 2009b, 2010). However, only by combining the psychological/sociological approach, based on a deeper insight into individual and group decisions and actions and the quantitative tools based on developments of statistical physics (for example studying the role of information and entropy in human communication) can one achieve reliable insights. Both approaches are important to improve the predictive capacity in the analysis of social opinions.

Psychology and sociology based research provides understanding of individual and group reactions, while statistical and simulation methods enable turning these into working models of social phenomena. Such a combined approach would be particularly applicable to the analysis of internet communication activities, as they are connected with large quantities of easily accessible data (Bordia, 1996; Chau & Xu, 2006; Ding & Liu, 2010; O'Connor, Balasubramanyan, Routledge, & Smith, 2010; Wojcieszak, 2008; Wojcieszak et al., 2010).

In addition to the sentiments and expressed opinions, which can be, as we suggest, monitored via datamining tools, Web-based communities provide direct access to social structure data, which is usually difficult to gather. This includes the social network topology and evolution (including the birth of new communities), temporal patterns of communication and group/individual activities (e.g. burstiness of communications and decay or persistence of reactions to events). These data may be compared with the simulations allowing improvements of model parameters and increased understanding of the underlying processes (see e.g. Barabasi, 2005; Schweitzer & Mach, 2008; Vázquez et al., 2006).

3.1. Opinion diffusion simulation

The two main approaches concerning the modeling of the diffusion of opinions are based on analogies from epidemiology and on various sub-models of interpersonal influence. The first approach is especially valuable in the case of opinions on a subject that is new to the population, what we could call *opinion formation* or *opinion spreading*. Such models assume for each agent three stages: the initial state, the alert state, and the percolated state. Such models draw on the works of Watts (Watts, 2002; Watts & Sheridan Dodds, 2007) in sociology and of Payne, Dodds, and Eppstein (Payne, Sheridan Dodds, & Eppstein, 2009) in physics and is here applied to the area of opinion cascading (Kaschesky & Riedl, 2011).

The other simulation approach focuses on *changes of opinions* under influence of other society members and/or external influences (media, propaganda). Here there are many detailed models describing the influences (such as the bounded confidence model or variants of the social impact model). Additionally more advanced models take into account the dynamic nature of the social interactions, due to which the network linking the citizens is not static, but evolves in parallel with their opinion changes, both influencing it and being influenced by it.

In addition, the behavior of agents should reflect individual differences between people. For example, MacKuen et al. (2010) introduce an interesting division of participants of political debates into the *deliberative citizen*, who considers – in a balanced way – all available information (including that opposing his/her current views) and the *partisan combatant* who is a passionate supporter of a single viewpoint. Such difference is relatively straightforward to include in agent-based computer simulations – and may crucially influence the spread of ideas and conflict over time. One of the most important non-deliberative factors influencing the attitudes toward specific subjects is emotion: emotions may facilitate or inhibit deliberation and, for example, completely freeze the opinions despite arguments and data pointing otherwise. Such situations are encountered frequently for issues of great importance (e.g. political attitudes driven by *ad persona* arguments, or issues where there is already a lot of emotion involved, such as attitudes to abortion, human-caused Climate Change, terrorism or racial issues).

When modeling opinion formation in modern societies, it is important to go beyond simple person-to-person interactions. In general, the model should include also influences of perceived in-group opinions, out-group pressure (and possible paradoxical contrarian roles Galam, 2007; Martins & Kuba, 2010), as well as effects of mass media. Mutz and Martin (2002), writing about traditional media (e.g. press, TV), predicted that in political environments their effects would surpass the person-to-person influences. This is even stronger

in online social networks, where participatory nature allows much stronger identification with selected attitudes.

Without the risks and responsibilities associated with face-to-face contacts. The users not only react to the perceived opinions, emotions and information, they effectively create their own environment, via selective attention, assortative grouping and reliance on content *created* within like-minded groups.

3.2. Opinion formation simulation

When we turn from opinion spreading to opinion changes (on both individual and societal levels), the models based on tools from statistical physics typically focus on global properties of the modeled system, such as the average opinion, time to achieve consensus or the number of different views within a society (Deffuant et al., 2000; Hegselmann & Krause, 2002; Kacperski & Hołyst, 1999; Nowak et al., 1990; Sznajd-Weron & Sznajd, 2000). In most works, people's actions and characteristics are simplified to the opinion itself (often taking a simple binary form of pro or contra specific issue) and to the social contact structure (e.g. network of contacts). Only a limited number of works consider possible interactions and correlations between opinions on related subjects (forming a complex 'worldview' of a person). The reason is that the mathematical models taken from statistical physics are not suitable to handle such complexity.

An improvement is offered by agent-based computer simulations, which allow integration of agent descriptions that are more detailed on the micro-level of individual behavior and therefore enable the combination of observations across several levels, from the individual micro-levels to the aggregated macro-levels. Such non-classical socio–economic modeling goes beyond simplified economic models because it takes into account several and multi-faceted characteristics of individuals, rather than one monolithic characteristic (e.g. utility maximization).

The proposed approach, developed in the course of our work, is motivated by three goals. The first was to keep the models applicable to real life social situations (Sobkowicz, 2009a), preferably giving concrete descriptive as well as predictive capabilities. Thus the model takes into account the social structure of the analyzed social situation (e.g. presence or absence of social links fixed by the environment, such as family or work connections Sobkowicz, 2009b); presence or absence of 'special status' individuals (such as leaders) (Sobkowicz, 2010); modes of communication (for example highly formalized communication in scientific research Sobkowicz, 2011 or dynamic discussions on the internet fora Chmiel et al., 2011a, 2011b; Sobkowicz & Sobkowicz, 2010).

The second goal was to keep the model parameters simple enough to allow mapping between sociological observations (e.g., observed communication network structures, attitude dynamics, political views), psychological factors as well as agent roles and characteristics. This goal is related to a larger one: the models should be *meaningful*, in the sense of providing insight and understanding of the observations (stressed by Epstein Epstein, 2008).

The third goal that we have in constructing the models is flexibility in taking into account previously unrecognized variables (especially when they can be mined from observations). The recent expansion of the opinion-related model to include emotions is an example of such flexibility (Chmiel et al., 2011a; Sobkowicz & Sobkowicz, 2011, 2012).

The current model takes into account these goals: it provides capability for the agents to cut the social links with those they disagree with and form new links with agents sharing the same opinion or to keep them. This makes the network structure dynamic. It is also possible to include some links that cannot be broken (e.g. family or work relationships).

With respect to opinions we include not only the proponents or opponents of a given view but also agents with no preferred opinion (neutral agents). This allows significant change from models where only committed agents are present, changing both the social network and opinion change dynamics. Appropriate real life examples of application of the simulation include political preferences and highly controversial opinions on topics such as abortion. Within the model, the strength of influence between agents decreases with their social separation, reflecting the fact that our opinions are swayed less by remote acquaintances or strangers than by the closest associates.

Secondly, the opinion of a given agent may be changed in reaction to perceived cumulative social opinion of others, corresponding to a properly averaged 'peer pressure' rather than the individual encounters. Many of the classical models have stressed the importance of such individual contacts in opinion changes of agents, but the constant background of perceived opinions, resulting from numerous encounters and information on opinions held by other members is relevant. In a way, this can be described as each agent continuously measuring and responding to the 'discomfort' due to difference between own opinion and properly averaged opinions of other agents. It also becomes possible to simulate propagandist efforts via simple parameterized factor of external influence.

The topic of media influence in social simulations requires particular attention. Traditional media (e.g. press, TV) are typically described as unidirectional influences: from mass media to readers/viewers. This situation is changing: in many countries the polarization of media can be observed as result of positive feedback cycle: as users are selective in their choice of the information sources, the media managers adjust their presentation to cater to the preferences of their target groups. Traditional impartialness gives way to specific views, which, in turn, increase the selectiveness of the readers (Wojcieszak, 2010).

The media polarization enhances the effects of social separation resulting in strong, socially undesirable effects. Wojcieszak (2010) points out the effects of perceiving *false consensus*, resulting from selective information processing. When both the direct social environment and media accessed by a person are effectively separated from the rest of society, the social structure may become fossilized, diminishing chances of achieving necessary consensus on important issues. All these effects are included in the general scheme of the opinion formation model presented in Fig. 1.

The situation is more complex when we analyze internet-based media, with active participation and lack of inhibitions that are present in face-to-face interactions. There are many environments (such as discussion boards related to blogs or news sites, social networks such as Facebook or Twitter) where users, emboldened by relative anonymity, actually actively jump on opponents to promote conflict. The role of such exchanges driven by disagreement is relatively weakly studied. While they break the false consensus limitations, they do not serve the goal of achieving consensus – perhaps the reverse: they strengthen the differences between supporting groups.

With increasing popularity of these internet 'battlegrounds', their role must be better understood and methods of including deliberation into the exchanges is of utmost importance. From the research perspective, the efforts must combine automated analyses of the actual content of the fora with modeling aimed at simulation of the role of moderators, graphical presentation of discussions, measures decreasing negative emotions. As there is no fixed social network (the internet users actually create their own environment), we develop a model for opinion formation simulations that take into account the dynamic, message-based nature of internet communication activities.

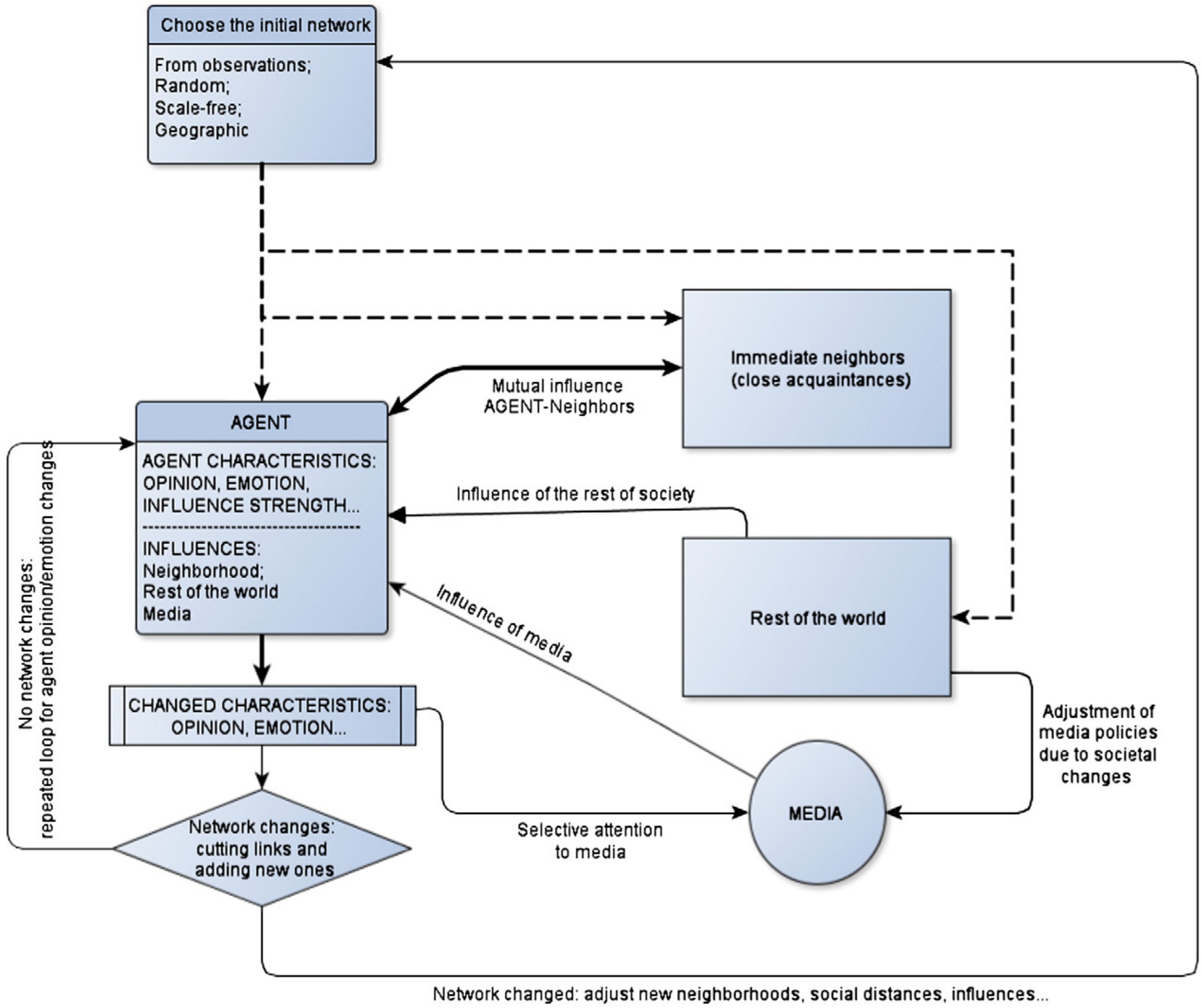

Fig. 1. Flowchart of agent-based for simulating opinion formation based on effects of leadership, dynamic social structure, presence of neutral agents and effects of social distance.

The process of creation of such message-based network (which may apply to discussion boards, blogs, e-mails etc.) within an agent-based model is schematically presented in Fig. 2.

One of the interesting aspects of internet-based communication is the broad range of observed behaviors. For example, the political blogs in the U.S. were shown to exhibit high levels of separation between the two opposing camps (Adamic & Glance, 2005). But there are environments where conflicted users actually seek contact with the other side and where most of the contacts are between the representatives of the opposing camps (Sobkowicz & Sobkowicz, 2011, 2012). The proposed model is able to simulate such vastly different situations, as shown in Figs. 3 and 4. Fig. 3 exhibits an environment where supporters of two opposing political views (gray and black) maintain intensive communication between the groups, while neutrals (white) play relatively small role, thus resembling a situation as observed in (Sobkowicz & Sobkowicz, 2011, 2012). Fig. 4 exhibits an environment where a separation between the majority and minority proponents (white and black) exists and neutrals (gray) provide the communication channel between the groups, thus resembling a situation as observed in (Adamic & Glance, 2005).

Simulation models based on complex agents, which include parameters for social distance, dynamic social network structure and nontrivial interactions, are gaining popularity and achieving increasing descriptive quality of real world descriptions. An example of the use of such models may be successful simulation of statistical and emotional properties of news media BBC discussions on politics and religion, comprised of millions of messages and thousands of distinct users (Chmiel et al., 2011a). More detailed description of such simulation is provided in the next section.

3.3. Opinion formation forecasting

Predictions within complex systems, such as communication networks, are possible thanks to statistical learning techniques. The analysis is often composed of three distinct modeling steps involving the prior knowledge about the opinions, the observation of the opinion diffusion and the posterior analysis (Watts, 2002): First, a model for the opinion diffusion network model is assumed. In a Bayesian framework, this corresponds to the definition of an a priori probability density function. The specific form of this prior distribution (i.e. the hyper-parameters) is taken from past observations, from experiences in Monte-Carlo simulations of the communication network and from the literature. Second, new observational data come with its own uncertainties (due to noise, partial information and error from the opinion mining system). The uncertainties are also represented by a probability distribution. Third, the observations and the prior opinion diffusion distribution are coupled together to update the distribution

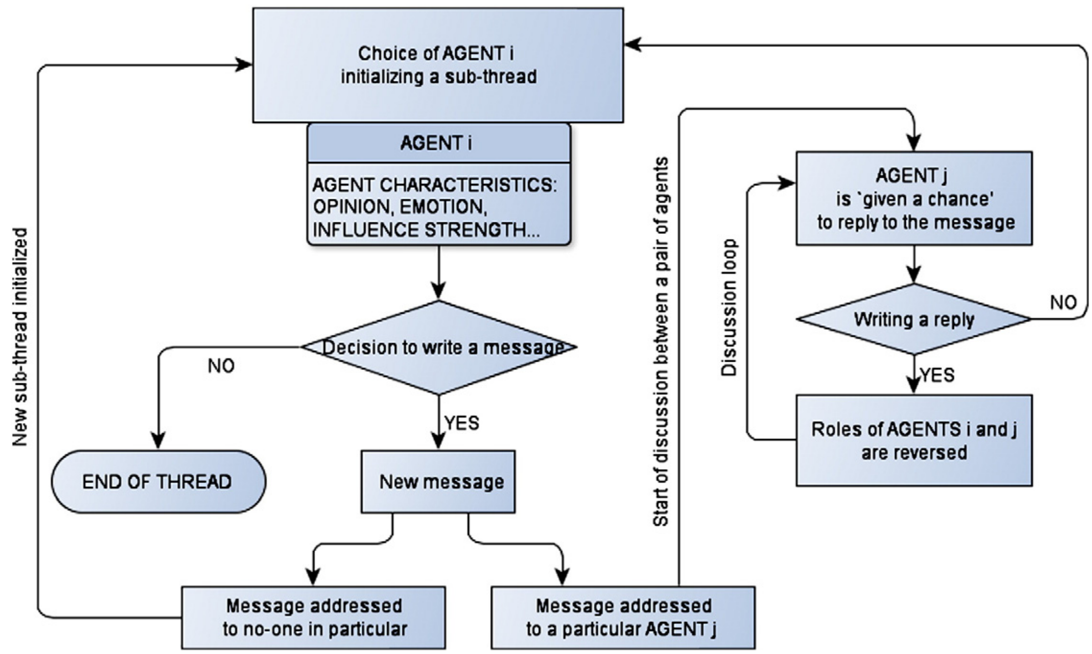

Fig. 2. Flowchart of agent-based model for simulating *ad hoc* networks, formed by message-based communication in the Internet, including person-to-person dialogues and messages addressed to the world at large.

and make a new network model consistent with the new data while following the prior assumption. This updated distribution is often called the *a posteriori* distribution in the Bayesian setting.

The new network model is used to make predictions and is a key component for interpretation: it defines the way the latent opinion diffusion process is mapped to the observations. But since in practice the opinion diffusion is only partially observed, no model is completely reliable, the uncertainty of the estimation has to be included in the analysis and the visualization of the results.

4. Application scenarios

The goal of this section is to provide a decision environment that presents the main historical and current developments regarding topics and opinions as well as trends of constituents' opinion in temporal and spatial (i.e. regional) contexts and the likely future evolution of the relevant communication networks in an intuitive and easy exercisable way. Using geospatial distributions of analytical results, decision makers understand the topics and opinions of different local, regional or global stakeholders based on their past opinions and sentiments towards a policy issue. The ultimate goal is to learn from unexpected reactions and the evolution of general, minority or viral opinions to bring forward accurate decisions and maximize the likelihood of intended consequences. The following paragraphs introduce some illustrative application scenarios and real-life examples to demonstrate possible uses.

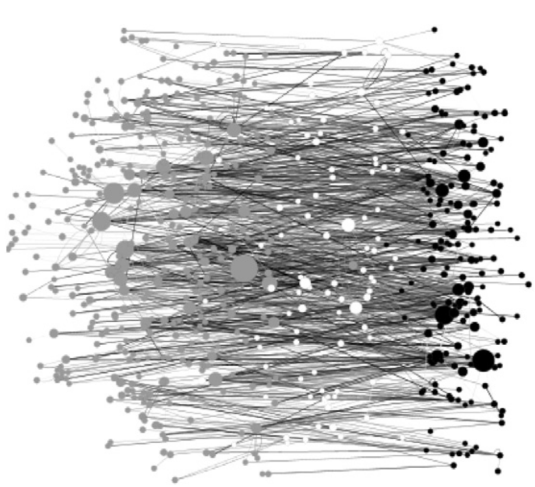

Fig. 3. Supporters of two opposing views (gray and black) maintain intensive communication between groups where neutrals (white) play relatively small role.

4.1. Political discussions in Poland

An important field of simulation-based analysis is highly polarized political environments, where support of active citizens is very stable divided in two almost equal camps. The first to apply sociophysics methods to such environments was Galam (2007). Stability of 50/50 political power division, observed in many countries, decreases chances of achieving general consensus in many important issues. The perception of diverse social questions through limited perspective of 'us vs. them' is especially dangerous when differences of views, initially limited to a few subjects, spill out to encompass all social issues. The resulting polarization of the society (including polarization of media) often leads to the situations where deliberative communication and working out of a consensus is almost impossible. Selective attention and false consensus (when one perceives only views of supporters of one's own view) lead to further extremism in the attitudes.

The contemporary political situation in Poland illustrates this very well. The two main parties are in the state of conflict since 2005, but the tragic crash of the plane carrying Polish president near Smolensk in April 2010 has resulted in an unprecedented split, both at the top politician level and among the citizens. For the supporters of the president killed in the crash, it was caused by terrorist attack, instigated by the opposing party. While initially, the mourning appeared to unite the nation, the discussions have, during the past two

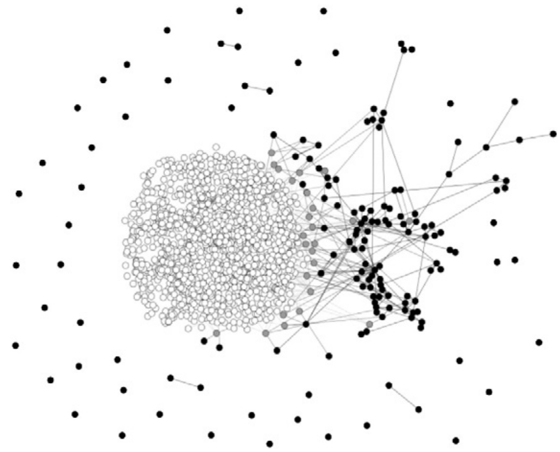

Fig. 4. Strong separation between the majority and minority proponents (white and black) where neutrals (gray) provide the communication channel between the groups.

years evolved into full scale attacks and accusations of national treason.

This polarization covers also mass media. Some TV and radio stations and newspapers *never* invite representatives of the opposing political camp. In 'real life' supporters of the two camps are very strongly separated (it is typical that whenever political differences occur, the matter is not discussed at work), which happens even at a family level. Interestingly, while some internet based communication channels (e.g. blogs) remain also separated, some, especially free-form discussion fora, serve as meeting grounds where supporters of opposing factions actively seek contact with the other side. Such environment provides extremely fertile ground to test agent based modeling techniques.

One can look for mechanisms that create network structure of communication (it is one of the very few social environments where social links may be recorded and observed *in situ*). With the advance of text recognition and lexical analysis tools, one can detect (to some extent) intentions, opinions and emotions expressed in the comments. These data mining tools, which are improving in quality may be used in two ways: *First*, they provide relatively cheaply the

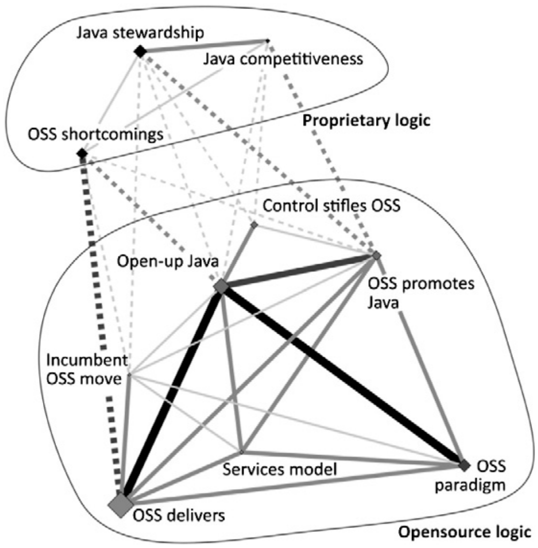

Fig. 5. Topics, centrality, momentum and cross-references of important issues in Phase 1.

information about the user behavior and characteristics. *Second*, they may serve as seeding the initial data for the simulation models. However, these models still require intensive 'training' and comparisons with estimates done by human evaluators. Using such techniques, we have performed a multi-year study of a highly polarized Polish political discussion forum (Sobkowicz & Sobkowicz, 2010, 2011), combining human and computer based analysis of content, goals and expressed emotions of several thousands of users, confirms the stability of individual user views over time.

The computer simulation closely reproduced many important aspects of the studied environment: communication network where most of the links were formed between political factions (the forum acted as a common 'battle ground' for the conflicted citizens, without the risks typically associated with face-to-face contacts). Simulations reproduced also the content statistics and details of emotion distribution for the posts, using simple and intuitive assumptions and few parameters (most of them taken from observed data). The same model, using data on evaluation that readers gave to the posted comments (thumbs-up and thumbs-down voting buttons) enabled the definition of the distribution of political views of forum *readers*, i.e. typically the 'invisible' 'read-only' part of the internet community. This distribution turned out to be the same as for the much smaller group of users writing comments (to within 1%). This illustrates that agent-based computer simulations can provide a window into data that is inaccessible by direct observation.

It is worth noting that these studies, focused on specific political environment in Poland characterized by deep polarization and high conflict level, appear to confirm Galam's suggestions of stability of bipolar division (Galam, 2007). These observations were confirmed by results of parliamentary elections in October 2011, where the two conflicted parties received almost the same support levels as four years earlier.

4.2. Governance of Java standard

The role of online opinions in the governance of the Java software standard illustrates the link between online opinion diffusion and its impact on policy making (in this case decisions of the Java Governing Board on opensourcing Java) (Kaschesky & Riedl, 2009). Decision-making on Java governance used to be a closed-book exercise involving the largest players. Small software firms and individual developers had to accept what the Java Governing Board decided. But this community with high internet-affinity informed and communicated via online media and forums to address this issue and request changes leading to opensource the Java standard.

Figs. 5 and 6 each present the two opposing communication networks on the issue at different points in time (Opensource logic vs. Proprietary logic). One communication pattern is around the 'proprietary logic' while the other is promoting the 'opensource logic'. On the left side, the communication network in 2002 is depicted, showing that Java opensource software was an issue that attracted some interest. On the right side, the same communication network is depicted in 2004, showing a massive increase in interest and engagement.

The proposed approach in this paper goes beyond these observations in three directions: 1) by tracking opinion formation in real-time, 2) by simulating the evolution of the communication networks (e.g. emergence of isolated minority opinions), and 3) by predicting its future evolution based on past observations and statistical learning techniques.

4.3. BP oil spill

Policymaking after the Deepwater Horizon oil spill serves as another example for an application scenario (Kaschesky & Riedl,

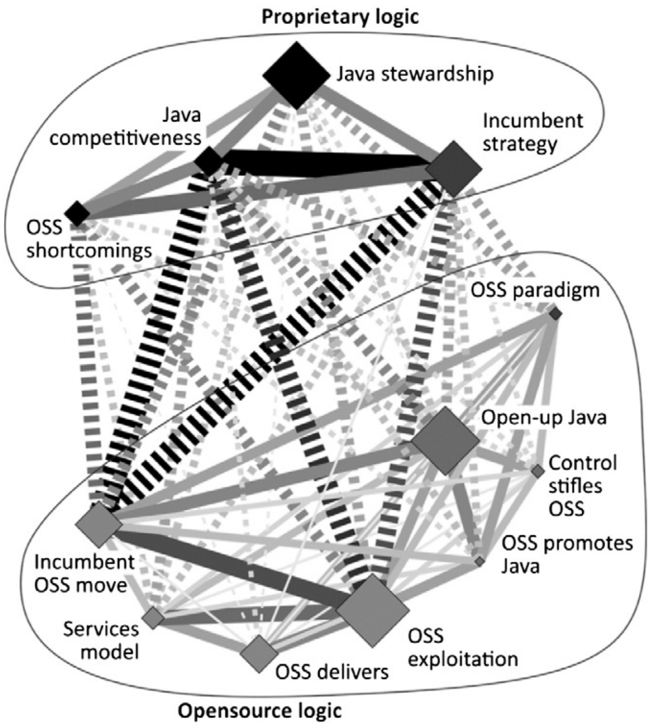

Fig. 6. Topics, centrality, momentum and cross-references of important issues in Phase 2.

2011). On 20 April 2010, an explosion on the oil rig caused by a blowout killed 11 crewmen and caused the second largest oil spill in history. Besides disaster relief operations, policymakers were reviewing the regulatory regime for oil exploration, the existing liability and compensation framework, the technological challenges involved with deepwater activities, and medium-term response activities (e.g. relief aid, use of chemical dispersants). At the same time, the general public accused BP and the government of inaction thereby asserting heavy pressure on policymakers to act swiftly. In addition, affected local citizens required medium-term help and support to cope with the consequences yet little was known to policymakers about the myriad of local problems that were caused by the oil spill.

Topic and opinion detection are illustrated using the example application scenario of the BP oil spill. Data collection retrieves online content related to the Deepwater Horizon oil spill. The field's boundaries are set so as to include all participants who exert some effect on opinion formation in the field.

For this example, the popular Technorati blog search engine is used for retrieving blog popularity rankings (arrows indicate popularity changes). Table 1 presents the 16 focal blogs representing the top 1 percent of all blogs related to U.S. Politics with an authority index above 1. Included in the sample are influentials or focal blogs, for example, the top 10 percent of blogs who maintain on average more connections to other blogs than do the remaining 90 percent. In addition, government news and publications and political news services such as Associated Press and Reuters are included for triangulation.

Topic detection may identify topics on the oil spill, such as 'Clean Energy Legislation', 'Nightmare Well', or 'oil spill' (uppercase words in Table 1). Opinion detection will then be able to analyze the content according to whether the topics are associated with primarily positive or negative opinions focusing on a specific region or the general public.

Let's take the post on the Greenpeace Campaign Blog which is used to illustrate opinion detection and sentiment analysis. In this context, the accumulation of words such as Tragedy, Accident, Pay

Table 1
Top focal blogs related to U.S. politics.

#	↑↓	Focal blog	Most recent post
1	→	Hot air	Quotes of the day
2	→	CNN political ticker	Congressman involved in on-camera confrontation
3	→	Think progress	Rep. Broun says CLEAN ENERGY LEGISLATION …
4	↓	Political punch	BP emails show disregard for 'NIGHTMARE WELL'
…		…	…
12	↓	RedState	TN state rep: You have to lift …
13	↓	TPMMuckraker	Gov't GEOLOGIST spoke of vast economic …
14	↓	Power Line	Speaking of gangster government
…		…	…
18	↑	Greenpeace campaign blog	Deepwater horizon disaster and OIL SPILL will impact …

the price, Damage (in uppercase below) signify a negative sentiment (Table 2).

5. Conclusion, implications and validation

5.1. Conclusion

The role of internet communications and communities and their influence on politics has received mixed reactions: From enthusiastic reactions (e.g. in the case of the 'Facebook revolution' interpretation of the Arab spring Chen, 2011; Zhuo, Wellman, & Yu, 2011) to more conservative evaluations pointing out the role of the traditional ways of social communication (Anderson, 2011; Stepanova, 2011). Some authors stress the *unexpectedness* of the usage of the tools designed with commercial and entertainment goals in mind in political environments (Lewiński & Mohammed, 2012). Still, while the exact role of modern computer media may be debated, there is no denial that for many people they have become the source of information, important influencer of emotions and the way to organize activities and to make decisions. This indicates the importance of *proactive* use of analyses of such media in policy making. Sourcing a wide range of views and concerns, which is being made possible by the proliferation of user-generated content across the web, appears to enhance the effectiveness of policymaking by providing insights that are typically difficult to obtain, such as hidden costs and risks, likely winners and losers, or differing cultural perspectives.

The prominence of the 'social web' and of user-generated content online has created a new situation for the interaction between policymakers and citizens. Policymakers did not have many indicators of citizen opinions available except for sporadic surveys, making precise assessments of the policy impact on constituents' life almost impossible and, consequently, inhibiting the possibility to react swiftly to emerging societal challenges. What most people felt and thought about policy measures and how this influenced their opinions and subsequent decisions was inaccessible – a policymaking black box.

Table 2
Illustration of opinion detection and sentiment analysis based on post from Greenpeace Campaign Blog.

The TRAGEDY we're witnessing right now is but the latest in a long line of OIL SPILLS, be they from pipelines, tankers, or exploratory drill rigs like the DEEPWATER HORIZON. Each ACCIDENT brings CONGRESSIONAL INQUIRIES, finger pointing, scathing editorials and PUBLIC OUTRAGE, yet we as a nation are no closer to weaning ourselves from oil than we were after any other big oil spill. So long as we remain dependent on oil we will continue to PAY THE PRICE IN HUMAN LIVES, as well as in ENVIRONMENTAL AND ECONOMIC DAMAGE.

Rather, online or offline surveys and consultations are undertaken at great costs and expenditure of time while highly valuable qualitative information on potential benefits and consequences is often available online, particularly regarding controversial issues that attract wide interest. When implemented, the proposed opinion mining and computer modeling approach allows valuable ideas and discussions to be collected and analyzed.

The approach supports inclusiveness in policymaking by ensuring that policymakers take more comprehensively into account the impact that proposed policy measures may have on different groups who are affected by the policy, such as businesses, families, older people, ethnic minorities etc. It enables policymakers to extend their understanding of how the policy may actual impact various stakeholders and to see its implementation from the citizen's point of view, thereby minimizing the likelihood of unintended consequences and strengthen the legitimacy of policy measures.

A more effective policy implementation and better identification of benefits and consequences may be achieved in two ways: (1) Sourcing and integrating expert and lay stakeholder views and opinions regarding the impact of a policy measure and (2) predicting the evolution of constituents' opinions to better adjust policy implementation and communication. In addition, modeling could lead to practical 'what if' analyses of the consequences and reactions to policies and to ways they are introduced to the society, leading to more effective policy implementation and communication through early recognition of problems and opposition possibly leading to extreme positions. When implemented, the approach supports policymakers in assessing and anticipating potential policy impacts on public opinion throughout the policy cycle:

- *Agenda setting*: opinion tracking provides policymakers with issue-specific, policy-focused, on-topic perspectives and sentiments about a concrete problem that requires policy action. In this way, policymakers are better able to understand the pros and cons as well as the expected benefits and consequences voiced by citizens regarding the problem.
- *Policy formulation*: opinion forecasting enables policymakers to assess and anticipate the sentiment and likely impact of proposed policy measures on constituents' opinion within a region. The opinion simulation not only performs sentiment mapping for congruence between constituents' opinion on the specific issue and the corresponding policy action. It also predicts how constituents' opinion may evolve further taking into account past evolution of sentiment on the issue and the new policy action. In this way, it enables policymakers to better anticipate the impact of proposed policy measures on constituents' opinion and adapt policy implementation and communication.
- *Implementation and evaluation*: opinion tracking recognizes issue-specific and policy-focused arguments and sentiments of opponents and proponents about a concrete problem while opinion simulation analyzes the actual impact of policy measures on constituents' opinion and predicts its further evolution. In this way, it enables policymakers to recognize and respond to the root cause and better communicate expected benefits and consequences of the policy.

5.2. Practical validation

In order to become useful for policymakers, the proposed model should be compared with results observed from other sources and 'facts'. Therefore the analytical and simulated results should be compared with surveys and polls and in user testing documenting the changes in social reactions to specific topics and policy implementations. Such comparisons extend in two directions:

- temporal, predicting the evolution of social trends; and
- representational, predicting the behavior of large social groups based on data from smaller samples.

Feedback from such test cases allows the opinion tracking and simulation to become more reliable over time, and provide direct measurement as to the quality of the analysis, modeling and simulation programming. By gradually improving the models, their control parameters (as well as the datamining techniques), the goal of better understanding the complex social processes may be brought nearer.

References

Adamic, L., & Glance, N. (2005). The political blogosphere and the 2004 US election: Divided they blog. *Proceedings of the 3rd international workshop on Link discovery* (pp. 36–43).

Allen, C., Machleit, K., Kleine, S., & Notani, A. (2005). A place for emotion in attitude models. *Journal of Business Research*, *58*, 494–499.

Anderson, L. (2011). Demystifying the Arab spring. *Foreign Affairs*, *90*, 2–7.

Banisch, S., Araujo, T., & Louçã, J. (2010). Opinion dynamics and communication networks. *Advances in Complex Systems*, *13*, 95–111.

Barabasi, A. (2005). The origin of bursts and heavy tails in human dynamics. *Nature*, *435*, 207–211.

Bordia, P. (1996). Studying verbal interaction on the internet: The case of rumor transmission research. *Behavior Research Methods, Instruments, & Computers*, *28*, 149–151.

Castellano, C., Fortunato, S., & Loreto, V. (2009). Statistical physics of social dynamics. *Reviews of Modern Physics*, *81*, 591–646.

Castells, M. (1996). *The rise of the network society*. Oxford: Blackwell Publishing.

Chau, M., & Xu, J. (2006). Mining communities and their relationships in blogs: A study of online hate groups. *International Journal of Human Computer Studies*, *65*, 57–70.

Chen, T. (2011). How networks changed the world. *IEEE Network*, *25*, 2–3.

Chesley, P., Vincent, B., Xu, L., & Srihari, R. (2006). Using verbs and adjectives to automatically classify blog sentiment. *Proceedings of AAAI-CAAW-06, the Spring Symposia on Computational Approaches.*

Chmiel, A., Sobkowicz, P., Sienkiewicz, J., Paltoglou, G., Buckley, K., Thelwall, M., & Holyst, J. (2011a). Negative emotions boost users activity at BBC Forum. *Physica A*, *390*, 2936–2944.

Chmiel, A., Sienkiewicz, J., Thelwall, M., Paltoglou, G., Buckley, K., Kappas, A., & Hołyst, J. (2011b). Collective emotions online and their influence on community life. *PloS one*, *6*, e22207.

Christakis, N., & Fowle, J. (2009). *Connected: The surprising power of our social networks and how they shape our lives – how your friends' friends' friends affect everything you feel, think, and do.* New York: Little Brown and Company.

Cialdini, R., Levy, A., Herman, C., Kozlowski, L., & Petty, R. (1976). Elastic shifts of opinion: Determinants of direction and durability. *Journal of Personality and Social Psychology*, *34*, 663.

Crano, W., & Prislin, R. (2006). Attitudes and persuasion. *Annual Review of Psychology*, *57*, 345–374.

Deffuant, G., Neau, D., Amblard, F., & Weisbuch, G. (2000). Mixing beliefs among interacting agents. *Advances in Complex Systems*, *3*, 87–98.

Derks, D., Fischer, A., & Bos, A. (2008). The role of emotion in computer-mediated communication: A review. *Computers in Human Behavior*, *24*, 766–785.

DeSteno, D., Petty, R., Rucker, D., Wegener, D., & Braverman, J. (2004). Discrete emotions and persuasion: The role of emotion-induced expectancies. *Journal of Personality and Social Psychology*, *86*, 43.

Ding, F., & Liu, Y. (2010). Modeling opinion interactions in a BBS community. *The European Physical Journal B*, *78*, 245–252.

Eaton, A., & Visser, P. (2008). Attitude importance: Understanding the causes and consequences of passionately held views. *Views Social and Personality Psychology Compass*, *2*, 1719–1736.

Epstein, J. M. (2008). Why model? *Journal of Artificial Societies and Social Simulation*, *11*(4), 12.

Friedkin, N., & Johnsen, E. (1999). Social influence networks and opinion change. *Advances in Group Processes*, *16*, 1–29.

Galam, S. (2007). From 2000 Bush–Gore to 2006 Italian elections: Voting at fifty-fifty and the contrarian effect. *Quality and Quantity*, *41*, 579–589.

Gastil, J., Black, L., & Moscovitz, K. (2008). Ideology, attitude change, and deliberation in small face-to-face groups. *Political Communication*, *25*, 23–46.

Griffiths, M. (2004). E-citizens: Blogging as democratic practice. *Electronic Journal of E-Government.*, *2*(3), 155–166.

Hegselmann, R., & Krause, U. (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artifical Societies and Social Simulation (JASSS)*, *5*(3).

Hołyst, J., Kacperski, K., & Schweitzer, F. (2001). Social impact models of opinion dynamics. *Annual Review of Computational Physics*, *20*, 531–535.

Huang, C., Tzou, P., & Sun, C. (2011). Collective opinion and attitude dynamics dependency on informational and normative social influences. *Simulation*, *87*, 875–892.

Inglehart, R., & Welzel, C. (2005). *Modernization, cultural change and democracy.* Cambridge UK: Cambridge University Press.

Kacperski, K., & Hołyst, J. A. (1999). Opinion formation model with strong leader and external impact: A mean field approach. *Physica A*, *269*, 511–526.

Kaschesky, M., & Riedl, R. (2009). Top-level decisions through public deliberation on the internet. *Proceedings of dg.o*.

Kaschesky, M., & Riedl, R. (2011). Tracing opinion-formation on political issues on the internet. *Proceedings of Hawaii international conference on system sciences*.

Kenny, C. (1994). The microenvironment of attitude change. *The Journal of Politics*, *56*, 715–728.

Kwon, N., Shulman, S. W., & Hovy, E. (2006). Multidimensional text analysis for eRulemaking. *Proceedings of dg.o*.

Lewiński, M., & Mohammed, D. (2012). Deliberate design or unintended consequences: The argumentative uses of Facebook during the Arab spring. *Journal of Public Deliberation*, *8*, 11.

Lin, W. H., Wilson, T., Wiebe, J., & Hauptmann, A. (2006). Which side are you on? Identifying perspectives at the document and sentence levels. *Conference on computational natural language learning*.

MacKuen, M., Wolak, J., Keele, L., & Marcus, G. E. (2010). Civic engagements: Resolute partisanship or reflective deliberation. *American Journal of Political Science*, *54*, 440–458.

Martins, A. C. R., & Kuba, C. D. (2010). The importance of disagreeing: Contrarians and extremism in the CODA model. *Advances in Complex Systems (ACS)*, *13*, 621–634.

Mitrović, M., Paltoglou, G., & Tadić, B. (2010). Networks and emotion-driven user communities at popular blogs. *The European Physical Journal B: Condensed Matter and Complex Systems*, *77*, 597–609.

Moss, S., & Edmonds, B. (2005). Towards good social science. *Journal of Artificial Societies and Social Simulation*, *8*(4), 13.

Mutz, D. C., & Martin, P. S. (2002). Facilitating communication across lines of political difference: The role of mass media. *American Political Science Review*, *95*, 97–114.

Nowak, A., & Lewenstein, M. (1996). Modeling social change with cellular automata. In R. Hegselmann, U. Mueller, & K. G. Troitzsch (Eds.), *Modelling and simulation in the social sciences from a philosophy of science point of view* (pp. 249–285). Dordrecht: Kluwer.

Nowak, A., Szamrej, J., & Latané, B. (1990). From private attitude to public opinion: A dynamic theory of social impact. *Psychological Review*, *97*(3), 362–376.

O'Connor, B., Balasubramanyan, R., Routledge, B., & Smith, N. (2010). From tweets to polls: Linking text sentiment to public opinion time series. *Proceedings of the international AAAI conference on weblogs and social media* (pp. 122–129).

Payne, J. L., Sheridan Dodds, P., & Eppstein, M. J. (2009). Information cascades on degree-correlated random networks. *Physical Review E*, *80*(2).

Petty, R., & Wegener, D. T. (1998). Attitude change: Multiple roles for persuasion variables. *The handbook of social psychology*, *1*. (pp. 323–390).

Petty, R., Wegener, D., & Fabrigar, L. (1997). Attitudes and attitude change. *Annual Review of Psychology*, *48*, 609–647.

Prabowo, R., & Thelwall, M. (2009). Sentiment analysis: A combined approach. *Journal of Infometrics*, *3*, 143–157.

Price, V., Cappella, J., & Nir, L. (2002). Does disagreement contribute to more deliberative opinion? *Political Communication*, *19*, 95–112.

Rainie, L., & Horrigan, J. (2007). Election 2006 online. *Pew internet & American life project report* (January).

Schweitzer, F., & Garcia, D. (2010). An agent-based model of collective emotions in online communities. *The European Physical Journal B: Condensed Matter and Complex Systems*, *77*, 533–545.

Schweitzer, F., & Mach, R. (2008). The epidemics of donations: Logistic growth and power-laws. *PLoS One*, 3.

Shulman, S., Hovy, E., Callan, J., & Zavestoski, S. (2006). Language processing technologies for electronic rulemaking. *Proceedings of dg.o*.

Skitka, L., & Sargis, E. (2006). The internet as psychological laboratory. *Annual Review of Psychology*, *57*, 529–555.

Sobkowicz, P. (2009a). Modelling opinion formation with physics tools: Call for closer link with reality. *Journal of Artificial Societies and Social Simulation*, *12*(1), 11.

Sobkowicz, P. (2009b). Studies of opinion stability for small dynamic networks with opportunistic agents. *International Journal of Modern Physics C (IJMPC)*, *20*(10), 1645–1662.

Sobkowicz, P. (2010). Effect of leader's strategy on opinion formation in networked societies with local interactions. *International Journal of Modern Physics C (IJMPC)*, *21*(6), 839–852.

Sobkowicz, P. (2011). Simulations of opinion changes in scientific communities. *Scientometrics*, *87*, 233–250.

Sobkowicz, P., & Sobkowicz, A. (2010). Dynamics of hate based internet user networks. *The European Physical Journal B: Condensed Matter and Complex Systems*, *73*, 633–643.

Sobkowicz, P., & Sobkowicz, A. (2011). Communication, emotions and opinions in political internet for a. *Cultural and opinion dynamics: Modeling, experiments and challenges for the future, ECCS2011 satellite meeting, Vienna*.

Sobkowicz, P., & Sobkowicz, A. (2012). Two year study of emotion and communication patterns in a highly polarized political discussion forum. *Social Science Computer Review*, http://dx.doi.org/10.1177/0894439312436512 (online first).

Stepanova, E. (2011). *The role of information communication technologies in the 'Arab Spring' implications beyond the region.* PONARS Eurasia Policy Memo no. 159.

Sznajd-Weron, K., & Sznajd, J. (2000). Opinion evolution in closed community. *International Journal of Modern Physics C: Computational Physics and Physical Computation*, *11*, 1157–1166.

Thelwall, M., Wilkinson, D., & Uppal, S. (2010). Data mining emotion in social network communication: Gender differences in MySpace. *Journal of the American Society for Information Science and Technology*, *61*, 190–199.

Theunis, M., Küster, D., Tsankova, E., & Kappas, A. (2010). Online discussion forums elicit subjective emotional response. *Third European conference on emotion, organized by the Consortium of European Research on Emotion (CERE), Villeneuve d'Ascq, France*.

Vázquez, A., Oliveira, J., Dezsö, Z., Goh, K., Kondor, I., & Barabási, A. (2006). Modeling bursts and heavy tails in human dynamics. *Physical Review E*, *73*, 036127.

Watts, D. J. (2002). A simple model of global cascades on random networks. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(9), 5766–5771.

Watts, D. J., & Sheridan Dodds, P. (2007). Influentials, networks, and public opinion formation. *The Journal of Consumer Research*, *34*(4), 441–458.

Wojcieszak, M. (2008). False consensus goes online. *Public Opinion Quarterly*, *72*, 781–791.

Wojcieszak, M. (2010). 'Don't talk to me': effects of ideologically homogeneous online groups and politically dissimilar offline ties on extremism. *New Media & Society*, *12*, 637–655.

Wojcieszak, M. E., Baeky, Y. M., & Delli Carpini, M. X. (2010). Deliberative and participatory democracy? Ideological strength and the processes leading from deliberation to political engagement. *International Journal of Public Opinion Research*, *22*.

Wood, W. (2000). Attitude change: Persuasion and social influence. *Annual Review of Psychology*, *51*, 539–570.

Zhuo, X., Wellman, B., & Yu, J. (2011). Egypt: The first internet revolt? *Peace Magazine*, *27*, 6–10.

Pawel Sobkowicz (Dr.) holds a PhD in solid state theory obtained at the Institute of Physics, Polish Academy of Sciences. He participated in the FP7 Cyberemotions project, analyzing the role of collective emotions in creating, forming and breaking e-communities (http://www.cyberemotions.eu). His focus is on modeling user interactions in web-based communications and on computational models of opinion formation.

Michael Kaschesky (Dr.) holds a PhD from Aston Business School, one of the leading management schools in Europe. Prior to that, he worked over 15 years in the multimedia and internet business. He publishes and presents regularly at academic and industry conferences. He is project coordinator of the FP7 Fusepool project implementing technology and customer intelligence for SMEs. At Bern University of Applied Sciences, he leads research projects to optimize the supply of timely, situation-aware, and needs-oriented information for decision support.

Guillaume Bouchard (Dr.) is a full time researcher in data mining in machine learning at XRCE. He received in 2001 an Engineering degree in applied mathematics from Institut National de Sciences Appliquées and a Master of Mathematics from the Université de Rouen. He joined XRCE in 2004 after receiving a PhD in Applied Mathematics from Institut National de Recherche en Informatique et en Automatique (INRIA) and Université Joseph Fourier, where he participated to the European project LAVA. Since then, his work mainly focused on theoretical analysis of pattern recognition algorithms and the development of statistical models for user profiling and causal models of failures in electronic devices. His current research interest focuses on artificial intelligence techniques to support content creation in knowledge management systems. He is also an active member of the PASCAL2 network of excellence since its creation.